

# Tight Logarithmic Asymptotic for the Probability of $n \times n$ Random Matrix with Uniform Distributed $\pm 1$ Entries to be Singular

Vladimir Blinovsky\*

Instituto de Matematica e Estatística, USP,  
Rua do Matao 1010, 05508-090, São Paulo, Brazil  
Institute for Information Transmission Problems,  
B. Karetnyi 19, Moscow, Russia,  
vblinovs@yandex.ru

## Abstract

We prove that probability  $P_n$  that  $n \times n$  random matrix with uniform distributed  $\pm 1$  entries is singular satisfies asymptotic inequality

$$P_n = 2^{-n+o(n)}.$$

Denote  $P_n$  the probability that  $n \times n$  random matrix with uniform distributed entries  $\pm 1$  is singular.

Obvious lower bound for the value  $P_n$  is the probability that two rows or columns of the matrix are linear dependent:

$$P_n \geq n^2 2^{1-n} (1 + o(1)).$$

Conjecture is that this lower bound is asymptotically tight. The history of the problem of determining upper bound for  $P_n$  started in 1967 when in [2] Komlós proved that  $P_n = o(1)$ . In 1993 in [1] Kahn, Komlós and

---

\*The author was supported by the São Paulo Research Foundation (FAPESP), Project no 2014/23368-6 and NUMEC/USP (Project MaCLinC/USP).

Szemerédi proved that  $P_n < (\alpha + o(1))^n$  for some  $\alpha < 1$  very closed to 1. Actually that work established many interesting ideas which also used later in improvements of this bound. First such improvement was made in [4] by Tao and Vu, who improve  $\alpha$  to 0.95 and in later work [5] to 0.75. Their improvement add additive combinatorics as ingredient in the proof and needs better estimation of  $|\cos \varphi|$ . This last bound was improved by Bourgain, Wood and Vu in [6] to  $\alpha = 1/\sqrt{2} \sim 0.70711$ . The main issue of this improvement is the estimation of  $\int_0^{2\pi} \prod_{i=1}^n \cos(\varphi m_i) d\varphi$  via Jensen inequality can help to find better bound, but as one can see from this work usage of Jensen inequality leads to loss in the estimations which can be avoided.

The main Theorem is the following

**Theorem 1** *The following inequality is valid*

$$P_n < 2^{-n+o(n)}.$$

The main idea behind the proof of this theorem is to refuse to use Halash-type argument which is important in the proof of the upper bound for  $P_n$  in [4], [5] and its modified version is used in [6] also. Define  $\mathcal{H} = \{x = (x_1, \dots, x_n) \in R^n : x_i = \pm 1\}$ . Define  $\Omega$  to be some universe set of integers of volume  $O(n^{n/2})$ , and for the hyperplane  $H \in R^n$ , with the property that there exists  $n - 1$  linear independent vectors from  $\mathcal{H}$  which belong to this hyperplane and the coordinates, which determine this hyperplane are integers from  $\Omega$ , the value  $P(H) = M(H)2^{-n}$  is the probability that some vector from  $\mathcal{H}$  belongs to  $H$  (we consider the uniform distribution on the set  $\mathcal{H}$ ). Here  $M(H) = |H \cap \mathcal{H}|$ .

Then as was shown in [3]

$$P(H) = \frac{1}{|\Omega|} \sum_{i \in \Omega} \prod_{j=1}^n \cos((\pi i a_j)/|\Omega|),$$

where  $(a_1, \dots, a_n)$  is the set of integers which are coordinates of hyperplane  $H$ . It was shown in [4] (Lemma 5.1) that

$$P_n = 2^{o(n)} \sum_H P(H). \quad (1)$$

The main issue of the work [1] is that the range of values of  $P(H)$  should be divided into three parts- small  $G_1$ , intermediate  $G_2$  and large  $G_3$ :

$$P_n = 2^{o(n)} \left( \sum_{H \in G_1} P(H) + \sum_{H \in G_2} P(H) + \sum_{H \in G_3} P(H) \right). \quad (2)$$

For small and large values of  $P(H)$  the method of estimating of the corresponding sum from (2) is the same: in [1] was proved the estimation

$$\sum_{H \in G_1} P(H) < n \max_{H \in G_1} P(H). \quad (3)$$

We next choose

$$\max_{H \in G_1} P(H) = 2^{-n+o(n)}.$$

For large  $P(H) = \Theta(n^{-1/2})$  in [1] was proved the estimation

$$\max_{H \in G_3} P(H) < n^2 2^{-n} (1 + o(1)).$$

Thus we need to estimate the value

$$\max_{H \in G_2} P(H),$$

where  $G_2 = \{H : P(H) \in [2^{-n+o(n)}, \Theta(n^{-1/2})]\}$ . We need the following key

**Lemma 1** *Let  $p_k$ ,  $k = \{-s, -s+1, \dots, -1, 0, 1, \dots, s-1, s\}$  be the symmetric probability distribution with  $p_0 = \frac{1}{2} \geq p_k$  and*

$$\begin{aligned} T(x, a) &= \sum_{j=0}^k p_k \cos((2\pi k x a)/|\Omega|), \\ T(H) &= \frac{1}{|\Omega|} \sum_{i \in \Omega} \prod_{j=1}^n T(i, a_i). \end{aligned}$$

*If for some constant  $C > 0$*

$$T(H)/C < P(H) < T(H)C \quad (4)$$

*is the range  $G_2$  of values of  $P(H)$ , then*

$$\sum_{H \in G_1} P(H) = 2^{-n+o(n)}.$$

Proof of this Theorem is essentially the same as the proof of Theorem from [5] (actually it is only part of that proof) and quite naturally use facts from additive combinatorics. We skip the proof and only mention that in order to establish upper inequality in (4) in [5] were used Halas-type argument and it

was before established the upper inequality for each  $\cos(\pi k a_i)$  in the product in the definition of  $T(H)$ . We suggest to go in another way and for some distribution  $p_k$  prove the upper inequality in (4) directly.

We consider the distribution  $p_k, k \in \{-1, 0, 1\}, p_0 = 1/2, p_{\pm 1} = 1/4$ . Then

$$T(x, a) = (1 + \cos((2\pi x a)/|\Omega|))/2$$

and

$$T(H) = \frac{1}{|\Omega|} \sum_{i \in \Omega} \prod_{j=1}^n (1 + \cos((2\pi i a_j)/|\Omega|))/2. \quad (5)$$

Define

$$f(i) = \prod_{j=1}^n ((1 + \cos((2\pi i a_j)/|\Omega|))/2)^{1/2}$$

$$\Lambda = \{i \in Z/QZ : f(i) \geq \epsilon\}.$$

We need the following fact, follows from Lemma 7.2 from [6]: for some  $C_1 > 1$

$$P(H)/C_1 < \frac{1}{|\Omega|} \sum_{i \in \Lambda} \prod_{j=1}^n (((1 + \cos((2\pi i a_j)/|\Omega|))/2)^{1/2} < P(H)C_1. \quad (6)$$

Thus

$$\begin{aligned} \frac{1}{|\Omega|} \sum_{i \in \Lambda} \prod_{j=1}^n (1 + \cos((2\pi i a_j)/|\Omega|))/2 &\geq \frac{\epsilon}{|\Omega|} \sum_{i \in \Lambda} \prod_{j=1}^n ((1 + \cos((2\pi i a_j)/|\Omega|))/2)^{1/2} \\ &\geq \epsilon P(H)/C_1 \end{aligned}$$

or for  $C = C_1/\epsilon$

$$CT(H) \geq \frac{C}{|\Omega|} \sum_{i \in \Lambda} \prod_{j=1}^n (1 + \cos((2\pi i a_j)/|\Omega|))/2 \geq P(H).$$

The following Lemma is the clone of Lemma 4.1 from [5]:

**Lemma 2** *If  $H \in G_2$  and*

$$T(H) > P(H)C,$$

*then*

$$\sum_{H \in G_2} P(H) < 2^{o(n)+n} C^{-\epsilon n},$$

*where  $\epsilon > 0$  can be chosen arbitrary small (actually we tend it to 0 sufficiently slowly and enlarge  $C$ , to satisfy last inequality with  $2^{-C_1 n}$  in the l.h.s with  $C_1 > 1$ ).*

Thus with high probability in  $G_2$  we have inequalities (4).

These considerations proved Theorem 1.

## References

- [1] J. Komlós, On the determinant of  $(0, 1)$  matrices, *Studia Sci. Math. Hungar.* 2 (1967) 7-22
- [2] J. Kahn, J. Komlós, E. Szemerédi, On the probability that a random  $\pm 1$  matrix is singular, *J. Amer. Math. Soc.* 8 (1995), 223-240.
- [3] J. Komlós, On the determinant of random matrices, *Studia Sci. Math. Hungar.* 3 (1968)
- [4] T. Tao and V. Vu, On random  $\pm 1$  matrices: Singularity and Determinant, *Random Structures and Algorithms* 28 (2006), no 1, 1-23.
- [5] T. Tao and V. Vu, On the singularity probability of random Bernoulli matrices, *J. Amer. Math. Soc.* 20 (2007) 603-628.
- [6] J. Bourgain, V. Vu, P. Wood, On the singularity of discrete random matrices, *Journal of Functional Analysis*, V.258, no. 2, (2010), 559-603